

GAGGLE SAFETY MANAGEMENT PROCEDURES

Gaggle's Safety Representatives review over a million blocked student communications each month. Due to such a high volume, the Gaggle Safety Management department has established many procedures and processes in order to ensure that reviewed student content is processed as efficiently and thoroughly as possible. This document aims to provide a basic explanation of the procedures that may be of greatest interest to customers and district Emergency Contacts.

Blocking Rules and Exceptions

Gaggle's machine learning algorithms are necessary, but they do not constitute a complete solution, which is why all blocked items are reviewed by Gaggle Safety Representatives. Very often, school-related projects contain harmless content that is flagged for inappropriate language or images. Examples include: artwork; projects on sexual education, drugs, and cyber-bullying; and reflection assignments on song lyrics and other types of popular media.

For this reason, Gaggle has established many exceptions to blocking rules, in which cases the student content will be designated as a "False Positive" and unblocked. All Safety Representatives are trained, so that they become acclimated with what content needs unblocked and what content needs to be brought to the attention of district Emergency Contacts.

Here are a few examples:

Flagged artwork is:

- + **Unblocked** if the student is making appropriate use of a painting or sculpture for an assignment. This includes pieces with nudity (e.g. the statue of David).
- + **Reported** if it is extremely violent, graphic, or if it includes photographs of nudity even if the photographs could potentially be labeled as art.

Images of guns are:

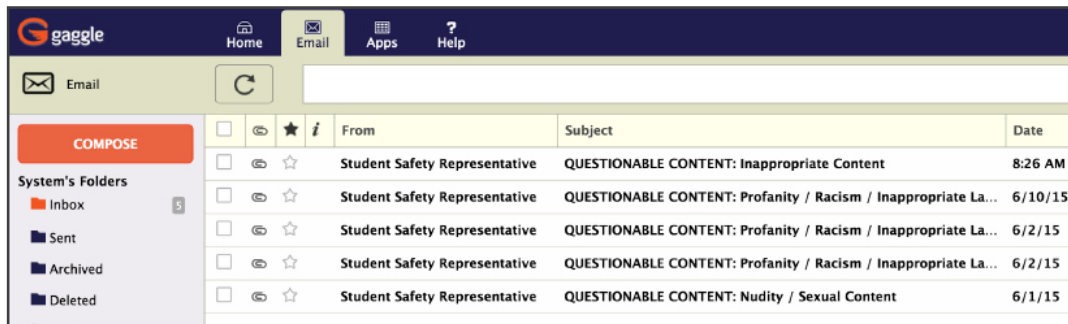
- + **Unblocked** if it is used appropriately for assignments.
- + **Unblocked** if it is in reference to hobbies, like hunting or paintball.
- + **Reported** if extremely graphic (e.g. assignment with Nazi officers with guns and victims).

Inappropriate and offensive song lyrics are:

- + **Deleted** if they are third party lyrics used just for fun.
- + **Reported** if the lyrics seem to be the student's own work for personal enjoyment.
- + **Reported** if homemade lyrics are excessively violent.
- + **Reported** if they are third party lyrics that are extremely graphic, but appear to be for an assignment.

Notification Categories

When student content is blocked and reviewed by Gaggle, it is assigned to one of many categories. These categories, then, display within the email subject line when Gaggle's Safety Representatives message school and district Emergency Contacts to report instances of Questionable Content, as seen in the screenshot below.



If you are a school or district Emergency Contact, see below for a short list of the categories you may encounter, as well as an explanation of what the category entails.

- + **Profanity/Racism/Inappropriate Language:** Instances of foul language, potential hate language, or other inappropriate language for the age level.
- + **Nudity/Sexual Content:** This can be communications between students, Victoria's Secret ads, adult dating sites, and other potentially harmful forms of advertisement. Usually district contacts are notified of student communications within this area, unless the student appears to be in immediate danger, in which case Safety Representatives will contact the district immediately. Spam or bulk mail is deleted.
- + **Violence/Abuse (self):** Cutting, suicide, or other self-harm language or threats. Usually these items would then be escalated or emergency alerted.
- + **Violence/Rape/Abuse (others):** Threats of violence to anyone or anything. This includes potential bullying situations.
- + **Alcohol/Smoking/Drugs:** Student reference to drug and alcohol use. Usually student communications within this category are escalated, unless the student appears to be in immediate danger, in which case Safety Representatives will contact the district immediately. Spam or bulk mail is deleted.
- + **Inappropriate Content:** Anything that does not fit into the other categories will fall within the inappropriate content category.